DNA Additives as a Mechanism for Unambiguously Marking Biological Samples

Inventors: Brian K. Haarer and Nigel S. Atkinson

Field of the Invention

The present invention relates generally to the forensic or medical analysis of biological samples. More specifically, the present invention relates to the marking of biological samples for their subsequent distinction from unmarked biological samples in standard forensic nucleic acid analysis.

Background of the Invention

In certain criminal investigations, "known" biological samples (primarily blood, saliva, and semen) are collected from victims, suspects, and their relatives. In these criminal investigations, "unknown" biological specimens are also collected, generally from the crime scene and from residences, automobiles, and other items associated with one or more suspects in the investigation. These unknown biological specimens are often scrapings of drops of blood, saliva, semen, or small tissue fragments.

Both known and unknown biological specimens are subjected to various analyses, including characterizations of their constituent deoxyribonucleic acid (DNA). The standard analysis methods used are (a) analysis of *Variable Numbers of Tandem Repeats (VNTR),* (b) analysis of *Short Tandem Repeats (STR),* (c) analysis of Single Nucleotide Polymorphisms (SNP), (d) analysis of *Restriction Fragment Length Polymorphisms (RFLPs),* and (e) analysis of mitochondrial DNA sequences. VNTR and STR analyses utilize simple or multiplex *Polymerase Chain Reaction (PCR)* technology; RFLP analysis utilizes restriction enzyme digestion of DNA followed by DNA hybridization techniques with labeled DNA probes; and mitochondrial DNA sequence analysis utilizes a combination of PCR technology and conventional dideoxy ("Sanger") sequencing in a process known as *cycle sequencing.*

Results from the above analyses are used to compare the known and unknown samples to determine any possible relationships between the samples.

However, problems may arise due to the deliberate or inadvertent contamination of

unknown biological samples by previously collected known biological samples, or by subsequent samples due to the confusion of samples (e.g. during analysis). In the example of criminal forensic analysis, such contamination could arise when blood from a victim is collected at a particular location, then transported to the residence of a suspect and subsequently released at the suspect's residence.

Thus a need exists for a mechanism whereby collected known biological samples would be unambiguously marked and identified at the time of collection. Then, if a marked sample should happen to contaminate another locale, the sample would be recognized as a contaminant upon subsequent analysis. This would also safeguard against the confusion of samples during analysis, preventing a "known" sample from being mistaken as a sample collected from a crime scene.

Summary of the Invention

It is therefore an object of the present invention to prevent confusion of known forensic or other biological samples with unknown samples, by adding a DNA molecule of defined composition to the known sample at the time of sample collection.

It is a further object of the present invention to have the added DNA molecule be detectable through standard methods of forensic analysis, without interfering with, or confusing the results of, such methods of analysis.

It is a further object of the present invention to decrease, if not eliminate, the possibility that collected biological samples could subsequently serve to contaminate some other person, place, or thing without being detected as an obvious contaminant.

It is a further object of the present invention to provide confirmation that known and unknown samples have not been confused during analysis, and to provide instant notification if such confusion does take place.

It is a further object of the present invention to improve the integrity of the data, such that analysis results are more likely to hold up in court.

It is a further object of the present invention to provide a marker for the known samples that shows up during the course of forensic analysis, and therefore does not require special detection methods beyond the inclusion of appropriate DNA primers or hybridization probes,

depending on the method of analysis.

The present invention accomplishes the above and other objectives by providing a mechanism for marking biological samples (blood, semen, saliva, etc.) that are to be used for subsequent nucleic acid analysis. The method involves adding a nucleic acid (DNA) molecule of known sequence to the biological sample at the time of sample collection. The method further utilizes primers specific to the complementary strands of the added DNA such that they will direct the synthesis of another DNA molecule of known length when used in a standard or multiplex polymerase chain reaction (PCR). This provides an unambiguous identifying label for the collected forensic or medical samples, including blood, semen, saliva, urine, tissue, and mixtures of bodily fluids. When used with the supplied primers or DNA probe(s), PCR or nucleic acid hybridization techniques will produce or recognize DNA fragments of predetermined size(s), preventing errant confusion of said samples with other forensic or medical samples that do not contain the aforementioned DNA additive.

Brief Description of the Drawings

Figure 1 illustrates the basic DNA additive and single primer set.

Figures 2A-2C illustrate alternative methods for generating two distinct fragments from PCR reactions(s).

Figure 2A illustrates a single template used with two alternative sets of primers.

Figure 2B illustrates two templates used with a single set of primers.

Figure 2C illustrates two templates used with two alternative sets of primers.

Figure 3 illustrates restriction sites at ends of DNA additive for use with RFLP analysis.

Detailed Description of the Invention

The present invention is a process whereby a defined and unique marker DNA is added to biological samples collected from known individuals by standard methods. Such tagging of collected samples effectively prevents their subsequent confusion with samples ("unknowns") that do not contain the defined marker DNA. Examples of unknown biological samples would be those collected at crime scenes or from crime victims, including but not limited to blood scrapings, hair, semen, saliva, blood, tissue scrapings, urine, mixtures of body fluids, etc.

The present invention is consistent with commonly used techniques for forensic analysis.

These include Short Tandem Repeat (STR), Variable Number of Tandem Repeats (VNTR), Single Nucleotide Polymorphisms (SNP), Restriction Fragment Length Polymorphism (RFLP), and mitochondrial sequencing analysis methods. These techniques are well known to those skilled in the art, and are described in the following references, each of which is incorporated herein by reference in its entirety.

(i) General reference for RFLP, VNTR, STR, and mitochondrial sequencing (the second reference also includes SNP, though not by that name):

Lincoln, P.J., and J. Thomson, eds. 1998. Forensic DNA Profiling Protocols. Humana Press, Inc.

Landegren, U., R. Kaiser, C.T. Caskey, and L. Hood. 1988. DNA diagnostics—molecular techniques and automation. Science. 242:229-237.

(ii) RFLP analysis:

Botstein, D., R.L. White, M. Skolnick, R.W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-331.

(iii) VNTR analysis:

Nakamura et al. 1987. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science. 235:1616-1622.

(iv) STR analysis:

Edwards, A., A. Civitello, H.A. Hammond, and C.T. Caskey. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am. J. Hum. Genet. 49:746-756.

Ricciardone et al. 1997. Multiplex systems for the amplification of short tandem repeat loci: evaluation of laser fluorescence detection. Biotechniques. 23:742-747.

(v) SNP analysis:

Nickerson, D.A., R. Kaiser, S. Lappin, J. Stewart, L. Hood, and U. Landegren. 1990. Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. Proc. Natl. Acad. Sci. USA. 87:8923-8927.

Nikiforov, T.T., R.B. Rendle, P. Goelet, Y.H. Rogers, M.G. Kotewicz, S. Anderson, G.L.

5

Trainor, and M.R. Knapp. 1994. Genetic bit analysis: a solid phase method for typing single nucleotide polymorphism. Nucleic Acids Res. 22:4167-4175.

Ross, P.L., K. Lee, and P. Belgrader. 1997. Discrimination of single-nucleotide polymorphisms in human DNA using peptide nucleic acid probes detected by MALDI-TOF mass spectrometry. Anal. Chem. 69:4197-4202.

(vi) General PCR analysis:

Mullis et al. 1986. Specific enzyme amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symp Quant. Biol. 51:263-273.

Mullis, K.B., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn GT, K.B. Mullis, and H.A. Erlich HA. 1988. Primer-directed enzymatic amplification on DNA with thermostable DNA polymerase. Science 239:487-491.

In a preferred embodiment of the present invention, the DNA additive is introduced irreversibly into the known samples at the time of collection. This can be accomplished by providing the DNA additive in the most commonly used collection vessels, including vacutainer tubes, FTA Blood Stain Collection Cards™ (or similar blood collection and preservation systems), saliva collection swabs, etc.

The present invention comprises a fragment or fragments of DNA of known sequence that are introduced into known samples. The DNA is provided either as an insert within a plasmid host that allows amplification in *E. coli*, or is added as a linear fragment. The length of the DNA(s) is such that it/they (a) provide PCR product(s) of known lengths when used with appropriate oligonucleotide primers in a PCR reaction in conjunction with either STR or VNTR analysis, (b) provide restriction fragments of known lengths that can be detected with defined nucleic acid probes when used in RFLP analysis, and (c) generate a unique known DNA sequence when used with the appropriate oligonucleotide sequencing primer(s) in conjunction with mitochondrial sequencing (in cycle sequencing reactions).

Variations on the exact nature of the added DNA are shown in Figures 1-3. **Figure 1** illustrates the basic form of the DNA additive, where the DNA fragment of known length and sequence is used in conjunction with primers complementary to its two ends, such that a PCR reaction utilizing these three components (DNA additive and two primers) will produce a PCR

product of defined length. This length should be of similar size to products generated by standard STR or VNTR primer sets. The DNA additive consists of DNA fragment of known sequence **102** inserted into a plasmid vector, (portions of the plasmid vector adjacent to the insert are shown **104**). A single primer set consists of first PCR primer **106** and second PCR primer **108** which are used to identify the presence of the DNA additive. In the PCR, first PCR primer **106** and second PCR primer **108** amplify, from the DNA additive, a DNA fragment of known size and sequence.

Since STR and VNTR analyses differ in the size range of PCR products generated, it may be necessary to develop DNA additives that are capable of generating at least two sizes of PCR products. This can be accomplished by one of the methods illustrated in Figures 2A-2C.

Figure 2A illustrates a single DNA additive that has binding sites for two different primer sets; the primer set used is determined by the type of forensic analysis being performed. First primer set is shown as first primer **202** and second primer **204**. Second primer set is shown as third primer **206** and fourth primer **208**. A DNA fragment of known sequence **210** is inserted into a plasmid vector, (portions of the plasmid vector adjacent to the insert are shown **212**). For example, the first primer **202** and second primer **204** would be used in a PCR reaction with the DNA additive to generate a DNA fragment of a size consistent with VNTR analysis, while the third primer **206** and fourth primer **208** would be used in a separate PCR reaction with the DNA additive to generate a DNA fragment of a size consistent with STR analysis, which examines smaller DNA fragments than does VNTR analysis.

Figure 2B illustrates a means by which two distinct DNA additives could be provided with a single primer set. Here, first DNA fragment of known sequence **214** is inserted into a plasmid vector, (portions of the plasmid vector adjacent to the insert are shown **216**) and second DNA fragment of known sequence **218** is inserted into a plasmid vector, (portions of the plasmid vector adjacent to the insert are also shown **220**). The primer set consists of first primer **222** and second primer **224**. In this situation, use of first primer **222** and second primer **224** in PCR reactions simultaneously on the two DNA additives would generate DNA fragments consistent with multiple methods of analysis. PCR from template **214** would generate a DNA fragment consistent with methods such as VNTR, that analyze longer DNA fragments, and PCR from

template **218** would generate a DNA fragment consistent with methods such as STR, that analyze shorter DNA fragments.

Figure 2C illustrates a means by which two distinct DNA additives could be provided with two distinct primer sets. Here, two templates are used with two alternative sets of primers. A first DNA fragment of known sequence **226** and a second DNA fragment of known sequence **228** are each inserted into a plasmid vector, (portions of the plasmid vector adjacent to the first insert **230** and the second insert **232** are shown). First primer set is shown as first primer **234** and second primer **236**. Second primer set is shown as third primer **238** and fourth primer **240**.

The DNA additives in Figure 2A and 2C are compatible with DNA sequencing methods as shown; the additive of Figure 2B is only compatible when an additional unique set of primers specific for one of the two added DNA molecules is used for the DNA sequence analysis.

Figure 3 illustrates the necessary features for use of the DNA additive in RFLP analysis; this design can easily be incorporated into any of the additive methods described in Figures 2A-2C. Known restriction enzyme site(s) are engineered at or near the ends of the DNA insert, and are used to release a DNA fragment of known length and sequence from the DNA additive. Enzyme one **302** and enzyme two **304** can represent the same or different restriction enzyme recognition sequences, and can represent single or multiple known restriction enzyme recognition sequences.

In a preferred embodiment, the DNA additive comprises one or more of the following features:

(a) the added DNA is stable for a length of time comparable to the shelf life of collected biological specimens;

(b) the added DNA and any primers used in conjunction with the added DNA do not interfere with the chosen form of analysis, and do not produce any PCR products, restriction fragments, bands detected by hybridization analysis, or DNA sequence other than expected for the added DNA. This requirement holds for the effects of oligonucleotide primers specific to the added DNA on DNA from the known biological sample, and for the effects of oligonucleotide primers normally used in forensic analysis methods on the DNA additive.

(c) the added DNA is compatible with, and stable through, standard DNA preparation

procedures specific for the type of collection vessel and for the specific forensic analysis protocol used;

(d) the concentration of the added DNA is such that it will be present in molar ratios similar to those of analysis targets in the known biological sample after preparation of the sample for analysis; and

(e) the DNA additive, or products generated from the DNA additive (e.g. by PCR), is compatible with commonly used methodologies for the subsequent analysis of such products, including, but not limited to: DNA hybridization analysis, agarose gel electrophoresis, polyacrylamide gel electrophoresis, capillary electrophoresis, and matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry.

In practice, the present invention comprises adding a nucleic acid of known sequence to standard collection vessels for forensic and medical biological samples; collecting the biological samples in the collection vessel containing the known nucleic acid additive; extracting the DNA from the biological sample; providing nucleic acid primers complementary to the known nucleic acid additive and analyzing the extracted DNA using standard assay techniques in conjunction with the provided primers. Standard assay techniques are known to those skilled in the art and include but are not limited to PCR-based analysis of short tandem repeats; PCR-based analysis of variable numbers of tandem repeats; DNA hybridization analysis of restriction fragment length polymorphisms; and the sequencing of mitochondrial DNA. In alternative embodiments the method can comprise: 1/ the use of a single primer set with a single DNA additive to produce a PCR product of defined length; 2/ the use of multiple primer sets with a single DNA additive to produce fragments of different sizes in separate PCR reactions, 3/ a single primer set used with multiple DNA additives to produce multiple fragments in a PCR reaction, and/or 4/ multiple primer sets used with multiple DNA additives to produce multiple fragments in a PCR reaction.

To support the process of the present invention, oligonucleotide primers or DNA hybridization probes necessary for the appropriate detection of the added DNA could be supplied either independently or as components of commonly used assay kits.

In practice, it is preferred that oligonucleotide primers or DNA hybridization probes used for the detection of the DNA additive are used on all samples analyzed, regardless of whether

they contain the DNA additive. Therefore, it is advantageous to supply necessary primers as standard components in STR and VNTR PCR reactions, Further, it is preferred that (cycle) sequencing using primers specific for the DNA additive be performed on all samples being analyzed for mitochondrial DNA sequence. Finally, it is preferred that DNA hybridization probes specific for the DNA additive be used to probe all samples being tested by RFLP analysis.

Any given set of PCR primers might support amplification of a DNA fragment from the genomes of a small percentage of the population that had not been identified during the design and testing of the DNA additive. This could give rise to false positives. In practice, this does not present much of a problem since the spuriously amplified fragment will most likely be a different size than a fragment amplified from the DNA additive. Furthermore, DNA sequence analysis could be used to confirm that the fragment arose as a result of amplification from the DNA additive.

In a preferred embodiment, the additive will be designed so that it contains binding sites for more than two PCR primers, thereby providing a simpler method to prevent the production of false positives. In this case, the presence of the DNA additive can be confirmed by performing multiple PCR reactions, each using a different set of primers. Each PCR reaction would generate an amplification product of known size and sequence from the DNA additive. Samples that contain the DNA additive would produce positive PCR reactions with all of the primer sets. The ability to confirm the presence of the additive by multiplex PCR greatly increases the robustness of the procedure.

In an alternative embodiment, different DNA additives are used for different individuals to provide further protection against mislabeling the samples. Here, if the DNA additive which was to have been used for Jane Doe is found in the sample labeled John Smith, then additional sampling of the possibly misidentified individuals would be indicated.

Having thus described the basic concept of the invention, it will be readily apparent to those skilled in the art that the foregoing detailed disclosure is intended to be presented by way of example only, and is not limiting. Various alterations, improvements and modifications will occur and are intended to those skilled in the art, but are not expressly stated herein. These modifications, alterations and improvements are intended to be suggested hereby, and within the

scope of the invention. Accordingly, the invention is limited only by the following claims and equivalents thereto.